

CODING METHOD FOR DIGITAL DATA

Publication number: JP10207489

Publication date: 1998-08-07

Inventor: FUJII OSAMU

Applicant: SHARP KK

Classification:

- international: H04B1/66; H04B1/66; (IPC1-7): G10L7/04; G10L9/18; H03M7/30

- european: H04B1/66M

Application number: JP19970009605 19970122

Priority number(s): JP19970009605 19970122

Also published as:



EP0855805 (A2)

US6138101 (A1)

EP0855805 (A3)

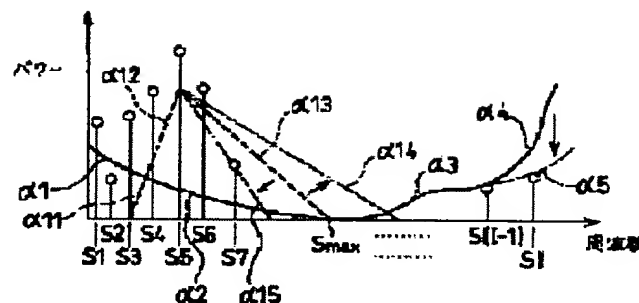
EP0855805 (B1)

DE69732619T (T2)

Report a data error here

Abstract of JP10207489

PROBLEM TO BE SOLVED: To enable selecting tone quality coinciding with hearing sensation of a listener by enabling changing the minimum listenable character previously obtained out of auditory sensation psychological characteristics without fixing to a characteristic of a person having general hearing sensation in auditory sensation psychological characteristics. **SOLUTION:** The minimum listenable character previously obtained out of auditory sensation psychological characteristics and a masking characteristic are made variable without fixing to a characteristic of a person having general hearing sensation in auditory sensation psychological characteristics. For example, a masked band is made wide and a masking level is made large by changing a curve of a masking characteristic from $\alpha 13$ to $\alpha 14$, bit allocation to a signal of which the power is relatively large is increased. In the same way, also in the minimum listenable characteristic, for example, a curve part shown in $\alpha 5$ is changed to a curve part shown in $\alpha 4$ by using weighting and offset in curve parts $\alpha 1$ - $\alpha 4$ of the minimum listenable characteristic based on auditory sensation psychological characteristic of a person having general auditory sensation characteristic. Thereby, bit allocation in a high frequency band is relatively increased.



Data supplied from the esp@cenet database - Worldwide

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平10-207489

(43) 公開日 平成10年(1998) 8月7日

(51) Int.Cl.⁶

識別記号

F I

G 1 0 L 7/04

G 1 0 L 7/04

G

9/18

9/18

A

H 0 3 M 7/30

H 0 3 M 7/30

A

審査請求 未請求 請求項の数 4 O L (全 12 頁)

(21) 出願番号

特願平9-9605

(22) 出願日

平成9年(1997) 1月22日

(71) 出願人 000005049

シャープ株式会社

大阪府大阪市阿倍野区長池町22番22号

(72) 発明者 藤井 修

大阪府大阪市阿倍野区長池町22番22号 シ

ャープ株式会社内

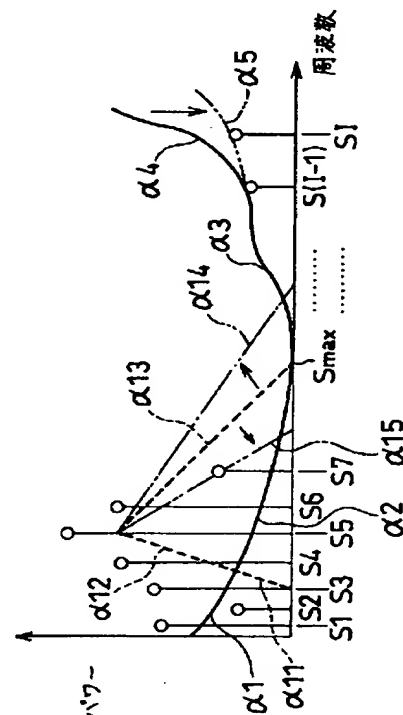
(74) 代理人 弁理士 原 謙三

(54) 【発明の名称】 デジタルデータの符号化方法

(57) 【要約】

【課題】 ミニディスクの圧縮符号化法である A T R A C で好適に用いられるデジタルデータの符号化方法において、聴取者の聴覚に一致した音質の選択を可能とする。

【解決手段】 一般的な聴覚を有する人の聴覚心理特性に基づいて作成される参照符 $\alpha 1$, $\alpha 2$, $\alpha 3$, $\alpha 4$ で示す最小可聴限特性および参照符 $\alpha 11$, $\alpha 12$, $\alpha 13$ で示すマスキング特性を、それぞれたとえば参照符 $\alpha 4 \rightarrow \alpha 5$ で示すように、また参照符 $\alpha 13 \rightarrow \alpha 14$ または $\alpha 15$ で示すように変更可能とすることによって、各周波数帯域へのビット割当てが変化し、聴取者の聴覚に一致した音質を選択することができるようになる。



【特許請求の範囲】

【請求項1】 楽音、音声等のデジタルデータを周波数領域に変換し、変換されたスペクトラムを複数の周波数帯域に分割し、各周波数帯域毎にビット割当てを行って符号化するにあたって、聴覚心理特性を反映して、前記各周波数帯域のパワーまたはエネルギーの大きさから各周波数帯域のマスキング閾値対雑音比を求め、該マスキング閾値対雑音比の大小に基づいて前記ビット割当てを行って符号化する方法において、

予め求められている最小可聴限特性を変更可能とすることを特徴とするデジタルデータの符号化方法。

【請求項2】 楽音、音声等のデジタルデータを周波数領域に変換し、変換されたスペクトラムを複数の周波数帯域に分割し、各周波数帯域毎にビット割当てを行って符号化するにあたって、聴覚心理特性を反映して、前記各周波数帯域のパワーまたはエネルギーの大きさから各周波数帯域のマスキング閾値対雑音比を求め、該マスキング閾値対雑音比の大小に基づいて前記ビット割当てを行って符号化する方法において、

前記パワーまたはエネルギーの大きさに対応して予め求められているマスキング特性を変更可能とすることを特徴とするデジタルデータの符号化方法。

【請求項3】 楽音、音声等のデジタルデータを周波数領域に変換し、変換されたスペクトラムを複数の周波数帯域に分割し、各周波数帯域毎にビット割当てを行って符号化する方法において、

聴覚心理特性を反映して、前記各周波数帯域のパワーまたはエネルギーの大きさから各周波数帯域のマスキング閾値対雑音比を求め、該マスキング閾値対雑音比の大小に基づいて前記ビット割当てを行う方法と、各周波数帯域のパワーまたはエネルギーの代表値に基づいてビット割当てを行う方法と、前記2種類の方法の処理にそれぞれ重み付けを行ってビット割当てを行う方法とを、切換え可能とすることを特徴とするデジタルデータの符号化方法。

【請求項4】 前記切換えを、各周波数帯域の相互に隣接するスペクトラムのパワーまたはエネルギーの差から求めたピークおよびローカルピークと、マスキング閾値との関係に対応して行うことを特徴とするデジタルデータの符号化方法。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】 本発明は、ミニディスクなどの記録媒体に楽音や音声等のデジタルデータを記録するにあたって、前記楽音や音声等に適応して各周波数帯域のスペクトルに対するビット割当てを行い、データ量を圧縮することができる符号化方法に関する。

【0002】

【従来の技術】 前記楽音や音声等のデジタルデータを高能率で圧縮符号化する方法として、前記ミニディスクで

用いられているATRAC(Adaptive Transform Acoustic Coding)が挙げられる。このATRACでは、高能率で圧縮するために、前記デジタルデータは、複数の周波数帯域に分割された後、可変長の単位時間でブロック化されてMDCT(Modified Discrete Cosine Transform)処理が施されてスペクトル信号に変換され、さらに聴覚心理特性を利用して割当てられたビット数で各スペクトル信号がそれぞれ符号化される。

【0003】 前記圧縮符号化に適用することができる前記聴覚心理特性には、等ラウドネス特性やマスキング効果が挙げられる。等ラウドネス特性は、同じ音圧レベルの音であっても、人間が感じ取る音の大きさが周波数によって変化することを表すものであり、したがって、人間が感じ取ることができる音の大きさである最小可聴限が、周波数によって変化することを表している。一方、マスキング効果には、同時マスキングと経時マスキングとがあり、同時マスキングは、複数の周波数成分の音が同時に発生しているときに、或る音が別の音を聴き取り難くさせる現象であり、経時マスキングは、大きな音の時間軸方向の前後では、マスキングを受ける現象である。

【0004】 前記経時マスキング効果を用いる例として、たとえば特開平5-91061号公報があり、この従来技術では、周波数変換のための単位時間内に過渡的信号を含む場合には、それ以前の単位時間の信号エネルギーまたはマスキング量に応じたワード長に基づいてビット割当てを決定することによって、プリエコーと呼ばれる音質劣化を防止している。また、たとえば特開平5-248972号公報では、マスキング閾値の決定に、過去の単位時間のスペクトラム分布を参照して経時マスキング効果を利用することによって、符号化効率を高める技術が提案されている。

【0005】 また、前記聴覚心理特性を利用した、たとえば反復法と呼ばれる割当て法では、入力デジタルデータに適応した実際のビット割当ては、以下のようにして行われている。まず、各周波数帯域のパワー S を求め、そのパワー S による他の周波数帯域に対するマスキング閾値 M を求める。次に、このマスキング閾値 M と、各周波数帯域を n ビットで量子化したときの量子化雑音パワー $N(n)$ とから、マスキング閾値対雑音比 MNR

$(n) = M / N(n)$ を求める。続いて、そのマスキング閾値対雑音比 $MNR(n)$ が最小となる周波数帯域にビットの割当てを行った後、該マスキング閾値対雑音比 $MNR(n)$ を更新し、再び最小の周波数帯域にビットの割当てが行われる。

【0006】

【発明が解決しようとする課題】 前記最小可聴限やマスキング閾値は、一般的な聴覚特性を備える人の聴覚特性がモデルとされている。したがって、聴取者の聴覚や好みの違いによって違和感を覚える場合がある。

3

【0007】たとえば、入力デジタルデータがホワイトノイズのように比較的スペクトラム成分がフラットな場合には、マスキング閾値は最小可聴限となってビット割当てが行われるので、ビットの多くは中低域に割当てられ、スペクトラム成分の大きさによっては、超低域と超高域とにビットが割当てられず、違和感を覚える場合がある。

【0008】また、入力デジタルデータが正弦波信号等のスペクトラム帯域の狭い信号と、前記ホワイトノイズとの合成波である場合には、正弦波信号の含まれる周波数帯域 f_1 だけがパワーが大きく、その周波数帯域 f_1 から遠い周波数帯域 f_2 程、パワーは急激に小さくなる。したがって、前記周波数帯域 f_2 では、正弦波信号によるマスキングの影響が殆どなく、自身のパワーによるマスキングの影響が一番大きくなる。これによって、前記周波数帯域 f_1 の信号対マスキング閾値の比 ($SMR = \text{自身のパワー} S \text{ とマスキング閾値} M \text{ との比}$) と、前記周波数帯域 f_2 の SMR との間には大きな違いが表れない。

【0009】すなわち、信号のパワーを S とし、各周波数帯域を n ビットで量子化したときの量子化雑音パワーを $N(n)$ とするとき、両者の相対的な関係から、マスキング閾値対雑音比 $MNR(n) = M/N(n) = (S/N(n)) / (S/M(n))$ は、前記周波数帯域 f_1 と f_2 とで同じような値となり、従来の適応ビット割当て方式では、前記マスキング閾値対雑音比 MNR

(n) のみに基づいてビット割当てを行っているので、周波数帯域 f_1 と f_2 とでビットの割当てが同じような値になってしまうという問題がある。

【0010】これによって、前記正弦波信号のマスキングの影響を受けない周波数帯域 f_2 が多数あると、該正弦波信号の含まれる周波数帯域 f_1 に配分されるビット数が相対的に小さくなり、該正弦波信号の量子化誤差が大きくなって音質が劣化するという問題もある。

【0011】この点、本件出願人は先に特開平7-202823号において、前記パワー S が小さい周波数帯域へのビット割当ての制限を自動的に行う構成を提案している。しかしながら、この従来の技術では、各周波数帯域に配分できるビットの上限値を、その周波数帯域のパワーに基づいて決定しているので、ホワイトノイズのパワーが大きいと、依然としてビット割当てに制限がかからない場合が生じるという問題がある。

【0012】本発明の第1の目的は、聴取者の聴覚に一致した音質を得ることができるデジタルデータの符号化方法を提供することである。

【0013】本発明の第2の目的は、スペクトラム帯域の狭い信号に対しても、音質劣化を防止することができるデジタルデータの符号化方法を提供することである。

【0014】

【課題を解決するための手段】請求項1の発明に係るデ

4

ジタルデータの符号化方法は、前記第1の目的を実現するために、楽音、音声等のデジタルデータを周波数領域に変換し、変換されたスペクトラムを複数の周波数帯域に分割し、各周波数帯域毎にビット割当てを行って符号化するにあたって、聴覚心理特性を反映して、前記各周波数帯域のパワーまたはエネルギーの大きさから各周波数帯域のマスキング閾値対雑音比を求め、該マスキング閾値対雑音比の大小に基づいて前記ビット割当てを行って符号化する方法において、予め求められている最小可聴限特性を変更可能とすることを特徴とする。

【0015】上記の構成によれば、聴覚心理特性が一般的な聴覚の人の特性に固定されておらず、該聴覚心理特性のうち、最小可聴限特性を変更可能とすることによって、不可聴帯域の小さいスペクトラム、または超低域や超高域のスペクトラムに対してビット割当てを行うか否かを選択することができるようになり、聴覚の優れた人や、個人の主観的な好みに対応し、聴取者の聴覚に一致した音質を得ることができる。

【0016】また、請求項2の発明に係るデジタルデータの符号化方法は、前記第1の目的を実現するために、楽音、音声等のデジタルデータを周波数領域に変換し、変換されたスペクトラムを複数の周波数帯域に分割し、各周波数帯域毎にビット割当てを行って符号化するにあたって、聴覚心理特性を反映して、前記各周波数帯域のパワーまたはエネルギーの大きさから各周波数帯域のマスキング閾値対雑音比を求め、該マスキング閾値対雑音比の大小に基づいて前記ビット割当てを行って符号化する方法において、前記パワーまたはエネルギーの大きさに対応して予め求められているマスキング特性を変更可能とすることを特徴とする。

【0017】上記の構成によれば、聴覚心理特性が一般的な聴覚の人の特性に固定されておらず、該聴覚心理特性のうち、マスキング特性を変更可能とするので、たとえば臨界帯域内でマスキングされてしまうスペクトラムに対してビット割当てを行うか否かを選択することができる。聴覚の優れた人や個人の主観的な好みに対応し、聴取者の聴覚に一致した音質を得ることができる。

【0018】さらにまた、請求項3の発明に係るデジタルデータの符号化方法は、前記第1の目的を実現するために、楽音、音声等のデジタルデータを周波数領域に変換し、変換されたスペクトラムを複数の周波数帯域に分割し、各周波数帯域毎にビット割当てを行って符号化する方法において、聴覚心理特性を反映して、前記各周波数帯域のパワーまたはエネルギーの大きさから各周波数帯域のマスキング閾値対雑音比を求め、該マスキング閾値対雑音比の大小に基づいて前記ビット割当てを行う方法と、各周波数帯域のパワーまたはエネルギーの代表値に基づいてビット割当てを行う方法と、前記2種類の方法の処理にそれぞれ重み付けを行ってビット割当てを行う方法とを、切換え可能とすることを特徴とする。

【0019】上記の構成によれば、ホワイトノイズなどの比較的スペクトラム成分がフラットなデータに対しては、周波数軸でフラットなビット割当てを行うことができ、また正弦波信号などの狭帯域の信号を含むデータに対しては、その狭帯域の信号への重点的なビット割当てを行うこともできるようになり、楽音のソースに適応した音質の選択が可能となる。

【0020】また、請求項4の発明に係るデジタルデータの符号化方法は、前記第2の目的を実現するために、前記請求項3におけるビット割当て法の切換えを、各周波数帯域の相互に隣接するスペクトラムのパワーまたはエネルギーの差から求めたピークおよびローカルピークと、マスキング閾値との関係に対応して行うことを特徴とする。

【0021】上記の構成によれば、前記ホワイトノイズなどのような広帯域なデジタルデータから、前記正弦波信号などの狭帯域のデジタルデータにまで適応して、最適なビット割当てを自動的に行うことができ、マスキング閾値対雑音比などの同時マスキングを利用したビット割当てに不向きな楽音に対しても、音質の劣化を防止することができる。

【0022】

【発明の実施の形態】本発明の実施の第1の形態について、図1～図3に基づいて説明すれば以下の通りである。

【0023】図1は本発明の実施の第1の形態の符号化方法を説明するための周波数スペクトラム図であり、図2は本発明の一適用例のミニディスク録音再生装置1の電気的構成を示すブロック図である。まず、図2を参照して、コンパクトディスク再生装置や、衛星放送受信装置などのデジタル音声信号源からは、入力端子2に、たとえば光信号でデジタルデータがシリアル入力される。前記光信号は、光電素子3において電気信号に変換された後、デジタルPLL回路4に入力される。デジタルPLL回路4は、前記デジタルデータから、クロックの抽出を行うとともに、サンプリング周波数および量子化ビット数に対応したマルチビットデータを再現する。前記マルチビットデータは、たとえばコンパクトディスクの44.1kHz、デジタルオーディオテープレコーダの48kHz、または衛星放送(Aモード)の32kHzなどの各種のサンプリング周波数から、周波数変換回路5において、ミニディスクの規格に対応した44.1kHzのマルチビットデータにサンプリングレートが変換された後、音声圧縮回路6に入力される。

【0024】音声圧縮回路6は、前記ATRA方式によって入力データの圧縮符号化を行い、その符号化された音声データは、ショックブルーフメモリコントローラ7を介して、信号処理回路8に入力される。前記ショックブルーフメモリコントローラ7に関連してショックブルーフメモリ9が設けられており、このショックブルー

フメモリ9は、音声圧縮回路6から出力される音声データの転送速度と、信号処理回路8に入力される音声データの転送速度との差を吸収するとともに、後述する再生時における振動等の外乱による再生信号の中断を補間し、音声データを保護するためのものである。

【0025】信号処理回路8は、エンコードおよびデコードとしての機能を備えており、前記音声データをシリアルの磁界変調信号にエンコードしてヘッド駆動回路11に与える。ヘッド駆動回路11は、記録ヘッド12を光磁気ディスク13上の所定の記録位置に移動させるとともに、前記磁界変調信号に対応した磁界を発生させる。このとき、光磁気ディスク13の前記所定の記録位置には、光ピックアップ21からレーザ光が照射されており、これによって磁界に対応した磁化パターンが光磁気ディスク13上に形成されてゆく。

【0026】一方、光磁気ディスク13からは、前記磁化パターンに対応したシリアル信号が光ピックアップ21によって再生され、該信号は高周波(RF)アンプ22で増幅された後、前記信号処理回路8に入力されて前記音声データにデコードされる。デコードされた音声データは、前記ショックブルーフメモリコントローラ7およびショックブルーフメモリ9によって前記外乱による影響が除去された後、音声伸長回路23に入力される。音声伸長回路23は、前記ATRA方式による圧縮符号化の逆変換処理を行い、フルビットのデジタル音声信号に復調を行う。復調されたデジタル音声信号は、デジタル/アナログ(D/A)変換回路24によってアナログ音声信号に変換された後、出力端子25から出力される。

【0027】前記高周波アンプ22で増幅されたシリアル信号はまた、サーボ回路31に入力されており、このサーボ回路31は、再生されたシリアル信号にตอบสนองして、ドライバ回路32を介してスピンドルモータ33の回転速度をフィードバック制御し、これによって所望とする線速度での再生が可能となり、また送りモータ34の回転速度をフィードバック制御し、これによって光ピックアップ21の光磁気ディスク13の半径方向に対する変移、すなわちトラッキングを制御することができ、また光ピックアップ21のフォーカシングをフィードバック制御する。

【0028】前記サーボ回路31、光ピックアップ21、高周波アンプ22、信号処理回路8およびドライバ回路32などは、電源ON/OFF回路35によって電力付勢される。また、この電源ON/OFF回路35の電源ON/OFF動作や、後述する信号処理動作などが、システムコントロールマイコン36によって集中管理されている。このシステムコントロールマイコン36に関連して、曲名入力や選曲操作などとともに、後述する音質調整動作が可能な入力操作手段37が設けられている。

【0029】上述のように構成されたミニディスク録音再生装置1において、図1および図3を参照して、前記音声圧縮回路6における前記ATRAC方式に従う本発明の実施の第1の形態のビット割当て法を以下に説明する。

【0030】前記ATRAC方式では、前記44.1kHzでサンプリングされた音声データは、所定の周波数帯域、すなわち0～5.5kHzのLo帯域、5.5～11kHzのMiddle帯域および11～22kHzのHi帯域に分割、および分割された各周波数帯域毎に所定の時間フレームに亘る音声データが前記MDCT処理によって周波数領域に変換され、変換されたMDCT係数が図1で示すようなi個の各周波数帯域のスペクトルパワー S_i ($i=1, 2, \dots, I$ 、たとえば $I=25$)に、さらに変換される。こうして得られた各スペクトルパワー S_i に対応して、図3で示すようなビット割当て処理が行われることになる。

【0031】音声圧縮回路6は、テーブルROM6aを内蔵しており、このテーブルROM6a内には、前記ATRAC方式に従うマスキング特性および最小可聴限特性が記憶されている。前記最小可聴限特性は、図1において参照符 $\alpha 1$ 、 $\alpha 2$ 、 $\alpha 3$ 、 $\alpha 4$ で示す曲線で表され、また、マスキング特性は前記スペクトルパワー S_i および各周波数帯域の臨界帯域幅などに対応して求められており、たとえば図1に示すようなパワー分布に対して、参照符 $\alpha 11$ 、 $\alpha 12$ 、 $\alpha 13$ で示すような曲線となる。これら参照符 $\alpha 1 \sim \alpha 4$ で示す最小可聴限特性および参照符 $\alpha 11 \sim \alpha 13$ で示すマスキング特性は、一般的な聴覚特性を有する人の聴覚心理特性に従って作成された固定値である。

【0032】本発明の実施の第1の形態では、この最小可聴限特性およびマスキング特性は、変更可能とされている。具体的には、たとえばマスキング特性の場合、スペクトルパワーが大きくなる程、かつ周波数が高くなる程、他の周波数帯域をマスキングする範囲は大きくなり、図1の場合にピークパワーであるスペクトルパワー S_5 に対して、該スペクトルパワー S_5 が影響を及ぼす範囲の上限値 S_{max} は、 $\alpha 13 \times (1 \pm \Sigma k)$ で求められる。ただし、 Σk は重み付けの係数値である。この変数 k を前記テーブルROM6aに予め複数種類格納しておき、システムコントロールマイコン36のレジスタ36aによってその変数 k を切換えることによって、前記マスキング特性の曲線 $\alpha 13$ を、参照符 $\alpha 14$ から参照符 $\alpha 15$ で示す範囲で切換えることができる。前記変数 k は、前記入力操作手段37を介して、聴取者によって設定可能である。

【0033】たとえばマスキング特性の曲線を参照符 $\alpha 13$ から $\alpha 14$ に変更することによって、マスキングされる帯域が広く、かつマスキングレベルが大きくなり、パワーの小さい信号へのビット割当てが減少、または無

くなってしまふ。したがって、相対的にパワーの大きい信号へのビット割当てが増加し、パワーの大きい信号のダイナミックレンジが増加する。これに対して、マスキング特性の曲線を参照符 $\alpha 13$ から $\alpha 15$ に変更することによって、パワーの小さい信号へのビット割当てが増加し、相対的にパワーの大きい信号へのビット割当てが減少する。したがって、周波数レンジを拡大することができる。前記重み付けのほかにも、前記マスキング特性の曲線 $\alpha 13$ にオフセットを持たせるようにしても、同様の効果を得ることができる。

【0034】同様に、最小可聴限特性も、一般的な聴覚特性を有する人の聴覚心理特性に基づく最小可聴限特性の曲線 $\alpha 1 \sim \alpha 4$ において、前記重み付けやオフセットを用いることによって、たとえば参照符 $\alpha 4$ で示す曲線部分を参照符 $\alpha 5$ で示すように変更する。これによって、高周波帯域でのビット割当てが相対的に増大する。

【0035】図3を参照して、まずステップp1において、前記MDCT処理によって得られた各周波数帯域に含まれるMDCT係数の2乗和から、前記スペクトルパワー S_i が計算される。ステップp2では、音声圧縮回路6は、前記システムコントロールマイコン36のレジスタ36aの値を読み出し、前記テーブルROM6a内に格納している、前記係数 k などのマスキング特性の変更パラメータを選択する。ステップp3でも、前記ステップp2と同様に、最小可聴限特性の変更パラメータが選択される。

【0036】ステップp4では、予め求められて前記テーブルROM6aに格納されている基準のマスキング特性および最小可聴限特性に対して、前記ステップp2、p3で選択されたパラメータに対応して変更が行われ、かつ2つの特性が合成されて最終のマスキング閾値が決定される。すなわち、変更された最小可聴限特性の曲線が参照符 $\alpha 1$ 、 $\alpha 2$ 、 $\alpha 3$ 、 $\alpha 5$ で表され、マスキング特性の曲線が参照符 $\alpha 11$ 、 $\alpha 12$ 、 $\alpha 14$ で示されるとき、合成して得られた最終のマスキング閾値の曲線は参照符 $\alpha 1$ 、 $\alpha 12$ 、 $\alpha 14$ 、 $\alpha 3$ 、 $\alpha 5$ で示すようになる。

【0037】ステップp5では、各周波数帯域のインデックスを前記 i とすると、前記ステップp1で求められたスペクトルパワー S_i と、ステップp4で求められた各周波数帯域のマスキング閾値 M_i との比 $SMR_i = S_i / M_i$ が、全ての周波数帯域に亘って計算される。この比 SMR_i は、対数グラフでは各周波数帯域のスペクトルパワー S_i のマスキング閾値 M_i を超えている部分の長さに対応する。

【0038】次にステップp6では、各周波数帯域の前記スペクトルパワー S_i を n ビットで量子化したとき、該スペクトルパワー S_i と量子化雑音パワー $N_i(n)$ との比 $SNR(n) = S_i / N_i(n)$ が求められる。この比 $SNR(n)$ は、統計的には信号の特性に

応じた定数となるので、統計処理によって予め求めておいてもよい。この比 $SNR_i(n)$ と前記比 SMR_i との比から、マスキング閾値と量子化雑音パワーとの比 $MNR_i(n) = SNR_i(n) / SMR_i$ が求められる。

【0039】ステップp7では、各周波数帯域の量子化ビット数が以下のようにして割当てられる。前記ビット数 n を0から大きくしてゆき、その都度、各周波数帯域のマスキング閾値と量子化雑音パワーとの比 MNR_i

(n)を計算し、その比 $MNR_i(n)$ が最小となる周波数帯域から順にビットを割当ててゆき、前記量子化ビット数 n を更新する度毎に、同様に比 $MNR_i(n)$ が最小となる周波数帯域にビットの割当てを行い、所定の割当て可能ビット数となるまで割当てを行うと、各周波数帯域の語長が決定されて出力される。すなわち、前記スペクトルパワー S_i の絶対値が、閾値 M_i を超えた部分の長さが最も長い周波数帯域から順次ビット割当てが行われることになる。

【0040】このようにして、図1で示すように聴取者の好みに対応して変更した聴覚心理特性のマスキング閾値となるようにビット割当てを行うことができ、たとえば最小可聴限特性を変更することによって、不可聴域にある小さいスペクトラム、または超低域や超高域のスペクトラムに対して、ビット割当てを行うか否かを選択することができるようになる。また、マスキング特性を変更することによって、各周波数帯域のパワーおよび周波数に対応した臨界帯域などによって決定されるマスキング特性を変更することができ、比較的大きなパワーのスペクトラムによってマスキングされてしまうスペクトラムに対して、ビット割当てを行うか否かを選択することができるようになる。このようにして、聴取者の聴覚に一致した音質を得ることができる。

【0041】本発明の実施の第2の形態について、図4に基づいて説明すれば以下の通りである。

【0042】図4は、本発明の実施の第2の形態のビット割当て法を説明するためのフローチャートである。注目すべきは、このビット割当て法では、前記マスキング閾値対量子化雑音比 $MNR_i(n)$ によるビット割当てと、各周波数帯域のパワーまたはエネルギーの代表値であるスペクトルパワー S_i を n ビットで量子化したときの量子化雑音パワー $SN_i(n)$ によるビット割当てとの比率を、任意に設定可能とすることである。したがって、前記音声圧縮回路6のテーブルROM6a内には、予め設定された前記比率 x が格納されており、入力操作手段37からの操作に应答して、システムコントロールマイコン36のレジスタ36aが前記比率 x を選択する。

【0043】まず、ステップp11では、前記ステップp1と同様に、前記MDCT係数の2乗和から、各周波数帯域のスペクトルパワー S_i が求められる。ステップp12では、システムコントロールマイコン36のレジ

スタ36a内の値が読出され、対応するテーブルROM6a内の比率 $x\%$ が選択される。

【0044】こうして決定された比率 x が0であるとき、すなわち1次割当て可能な総ビット数 B_1 が0であるときには、前記マスキング閾値対量子化雑音比 $MNR_i(n)$ によるビット割当てを行わずに、直接、後述するステップp18へ移る。これに対して、前記総ビット数 B_1 が0でないときには、ステップp13へ移る。

【0045】ステップp13では、ミニディスクのオーディオスペクトラムデータの総ビット数 B_0 である1,144~1,464ビットに対して、前記比 $MNR_i(n)$ による1次割当ての可能な総ビット数 B_1 が、 $B_1 = B_0 \times x / 100$ から求められる。

【0046】ステップp14では、予め求められている一般的な聴覚を有する人の聴覚心理特性に対応した最小可聴限特性およびマスキング特性に対応したマスキング閾値、すなわち前記図1における曲線 α_1 , α_{12} , α_{13} , α_3 , α_4 が求められる。続いてステップp15, p16では、前述のステップp5, p6と同様にして、各周波数帯域のスペクトルパワー S_i とマスキング閾値 M_i との比 SMR_i から、マスキング閾値と量子化雑音パワーとの比 $MNR_i(n)$ が求められる。ステップp17では、前記ステップp7と同様にビット割当てが行われるけれども、このときの割当て総ビット数は、前記ステップp12で求められた1次割当て可能な総ビット数 B_1 である。

【0047】ステップp18では、前記量子化雑音パワー $SN_i(n)$ が求められ、ステップp19でその求められた量子化雑音パワー $SN_i(n)$ が最大となる周波数帯域にビット割当てが行われ、以降、順次、量子化雑音パワー $SN_i(n)$ が更新されてビット割当てが行われてゆき、2次割当て総ビット数 $B_2 = B_0(1 - x / 100)$ が無くなるまでビット割当てが行われる。なお、これらステップp18, p19では、前記1次割当て可能な総ビット数 B_1 が0であって、前記ステップp14~p17を行わずに、直接、ステップp13から該ステップp18, p19に移るとき、および $x = 100$ でないときには処理が行われ、 $x = 100$ 、すなわち $B_2 = 0$ であるときには、前記ステップp17から、直接、語長出力へと移る。

【0048】このようにして、ビット割当ての比率 x を変化可能とすることによって、入力信号が前述の正弦波信号とホワイトノイズとの合成波である場合のような、ローカルピークがなく、単一正弦波に類似した、たとえばソロのピアノ演奏である場合には、マスキング閾値対量子化雑音比 $MNR_i(n)$ のみでビット割当てを行ってしまうと、多くのビットがパワーの小さい雑音成分に割当てられてしまい、相対的にピアノの量子化誤差が大きくなってしまふのに対して、量子化雑音パワー $SN_i(n)$ に基づいてビット割当てを行うことによって、ピ

アノへのビット割当てを増やすことができ、ピアノの量子化誤差を低減することができる。

【0049】また、入力信号がたとえばオーケストラの演奏のように、多数のローカルピークを有する音と雑音とで構成される楽音である場合には、大きい信号の近くの帯域にある小さいローカルピークの成分の楽音や雑音をマスクしてビット割当てを無くし、マスクされない大きい信号へとビット割当てを行うことができる前記マスキング閾値対量子化雑音比 $MNR_i(n)$ に基づいてビット割当てを行うことによって、忠実度の高い録音を行うことができる。

【0050】さらにまた、入力信号がソロのクラリネット演奏などのように、3～4個程度の適度のローカルピークを有する楽音とノイズとの場合のように、先の2例の中間的な楽音である場合には、前記マスキング閾値対量子化雑音比 $MNR_i(n)$ と、量子化雑音パワー $SN_i(n)$ とに重み付けしてビット割当てを行うことによって、前記クラリネットの忠実度を向上することができる。

【0051】このようにして、種々の楽音ソースに最も好ましいビット割当て法を選択することができるようになる。

【0052】本発明の実施の第3の形態について、図5

$$M/(NM+1) = 0$$

したがってマスクされているローカルピークがない場合には、前記比率 x が0%となって前記1次割当て可能な

$$0 < M/(NM+1) \leq 0.5$$

であれば $x = 50 \sim 90\%$ でセットされ、

$$0.5 < M/(NM+1)$$

であれば $x = 100\%$ 、すなわちビット総数 B_0 の全てが1次割当て可能な総ビット数 B_1 となる。

【0056】前記ローカルピークの検出は、たとえば図6では、スペクトルパワーのピーク値 S_5 が検出されると、その周波数帯域を含む前後に所定の周波数帯域、図6の例では低周波帯域側に2つ、高周波帯域側に4つに亘って、それらの各周波数帯域におけるスペクトルパワー $S_3 \sim S_9$ 相互間の差 D_{34} , D_{45} , ..., D_{89} が極性とともにも求められ、その極性の変化および絶対値からローカルピークを検出することができる。

【0057】したがって、ビット割当て処理では、前記ステップp11およびステップp1と同様のスペクトルパワー S_i の計算がステップp21で行われると、ステップp22ではピーク値が求められ、ステップp23でそのピーク値によるマスキング特性を含めたマスキング閾値が求められる。ステップp24では、前記式1～3による判定結果から前記比率 x が求められ、1次割当て可能な総ビット数 B_1 が決定される。その後、ステップp25～p27では、前記ステップp15～p17と同様に、マスキング閾値対量子化雑音比 $MNR_i(n)$ によるビットの1次割当てが計算され、その後ステップp

および図6に基づいて説明すれば以下の通りである。

【0053】図5は、本発明の実施の第3の形態のビット割当て法を説明するためのフローチャートである。注目すべきは、このビット割当て法では、スペクトルパワー S_i のピークおよびローカルピークと、マスキング閾値との関係に適應して、前記マスキング閾値対量子化雑音比 $MNR_i(n)$ によるビット割当てと、量子化雑音パワー $SN_i(n)$ によるビット割当てとの前記比率 x を自動的に決定するようにしたことである。したがって、前記比率 x は以下のようにして決定される。

【0054】まず、全ての周波数帯域のスペクトルパワー $S_1 \sim S_I$ に亘って、図6において参照符 S_5 で示すようなピーク値が求められる。次に、そのピーク値によるマスキング特性を含めた、図6に示すようなマスキング閾値が求められる。続いて、各周波数帯域において、参照符 S_8 で示すようなローカルピークを検出し、そのローカルピークのうち、前記マスキング閾値でマスクされてしまう個数とマスクされない個数とを求め、それらの比に対応して前記比率 x が決定される。

【0055】すなわち、ローカルピークの総数を NM とし、マスクされているローカルピークの総数を M とするとき、

$$\dots (1)$$

総ビット数 B_1 が0にセットされる。これに対して、

$$\dots (2)$$

$$\dots (3)$$

28, p29では、前記ステップp18, p19と同様に、量子化雑音パワー $SN_i(n)$ によるビットの2次割当てが行われる。

【0058】これによって、前述の図4で示すような楽音に適應した音質の優れたビット割当てを、自動的に行うことができ、マスキング閾値対量子化雑音比 $MNR_i(n)$ によるビット割当てが不向きな楽音に対しても、音質の劣化を防止することができる。

【0059】なお、前記図1および図3で示すような最小可聴限特性およびマスキング特性の変更を、前記図4または図5で示すビット割当て法にさらに適用するようにしてもよい。

【0060】

【発明の効果】請求項1の発明に係るデジタルデータの符号化方法は、以上のように、楽音、音声等のデジタルデータを周波数領域に変換し、変換されたスペクトラムを複数の周波数帯域に分割し、聴覚心理特性を反映して各周波数帯域にビット割当てを行って符号化するにあたって、聴覚心理特性を一般的な聴覚の人の特性に固定するのではなく、該聴覚心理特性のうち、最小可聴限特性を変更可能とする。

【0061】それゆえ、不可聴帯域の小さいスペクトラム、または超低域や超高域のスペクトラムに対してビット割当てを行うか否かを選択することができるようになり、聴覚の優れた人や、個人の主観的な好みに対応し、聴取者の聴覚に一致した音質を得ることができる。

【0062】また、請求項2の発明に係るデジタルデータの符号化方法は、以上のように、楽音、音声等のデジタルデータを周波数領域に変換し、変換されたスペクトラムを複数の周波数帯域に分割し、聴覚心理特性を反映して各周波数帯域にビット割当てを行って符号化するにあたって、聴覚心理特性を一般的な聴覚の人の特性に固定するのではなく、該聴覚心理特性のうち、マスキング特性を変更可能とする。

【0063】それゆえ、たとえば臨界帯域内でマスキングされてしまうスペクトラムに対してビット割当てを行うか否かを選択することができ、聴覚の優れた人や個人の主観的な好みに対応し、聴取者の聴覚に一致した音質を得ることができる。

【0064】さらにまた、請求項3の発明に係るデジタルデータの符号化方法は、以上のように、楽音、音声等のデジタルデータを周波数領域に変換し、変換されたスペクトラムを複数の周波数帯域に分割し、聴覚心理特性を反映して各周波数帯域にビット割当てを行って符号化するにあたって、各周波数帯域のマスキング閾値対雑音比を求め、該マスキング閾値対雑音比の大小に基づいて前記ビット割当てを行う方法と、各周波数帯域のパワーまたはエネルギーの代表値に基づいてビット割当てを行う方法と、前記2種類の方法の処理にそれぞれ重み付けを行ってビット割当てを行う方法とを、切換え可能とする。

【0065】それゆえ、ホワイトノイズなどの比較的スペクトラム成分がフラットなデータに対しては、周波数軸でフラットなビット割当てを行うことができ、また正弦波信号などの狭帯域の信号を含むデータに対しては、その狭帯域の信号への重点的なビット割当てを行うこともできるようになり、楽音のソースに適応した音質の選択が可能となる。

【0066】また、請求項4の発明に係るデジタルデータの符号化方法は、以上のように、前記請求項3におけるビット割当て法の切換えを、各周波数帯域の相互に隣接するスペクトラムのパワーまたはエネルギーの差から

求めたピークおよびローカルピークと、マスキング閾値との関係に対応して行う。

【0067】それゆえ、前記ホワイトノイズなどのような広帯域なデジタルデータから、前記正弦波信号などの狭帯域のデジタルデータにまで適応して、最適なビット割当てを自動的に行うことができ、マスキング閾値対雑音比などの同時マスキングを利用したビット割当てに不向きな楽音に対しても、音質の劣化を防止することができる。

10 【図面の簡単な説明】

【図1】本発明の実施の第1の形態の符号化方法を説明するための周波数スペクトラム図である。

【図2】本発明の一適用例であるミニディスク録音再生装置の電氣的構成を示すブロック図である。

【図3】本発明の実施の第1の形態のビット割当て法を説明するためのフローチャートである。

【図4】本発明の実施の第2の形態のビット割当て法を説明するためのフローチャートである。

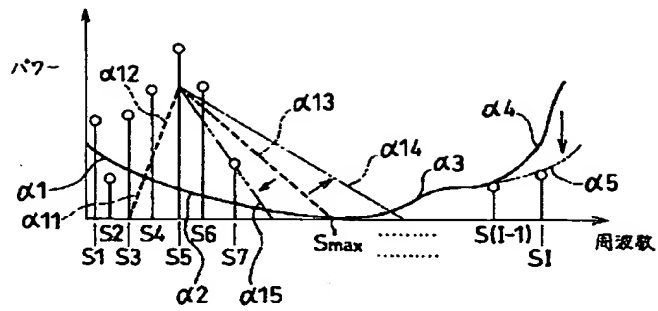
20 【図5】本発明の実施の第3の形態のビット割当て法を説明するためのフローチャートである。

【図6】図5で示すビット割当て法におけるピークおよびローカルピークの検出動作を説明するための周波数スペクトラム図である。

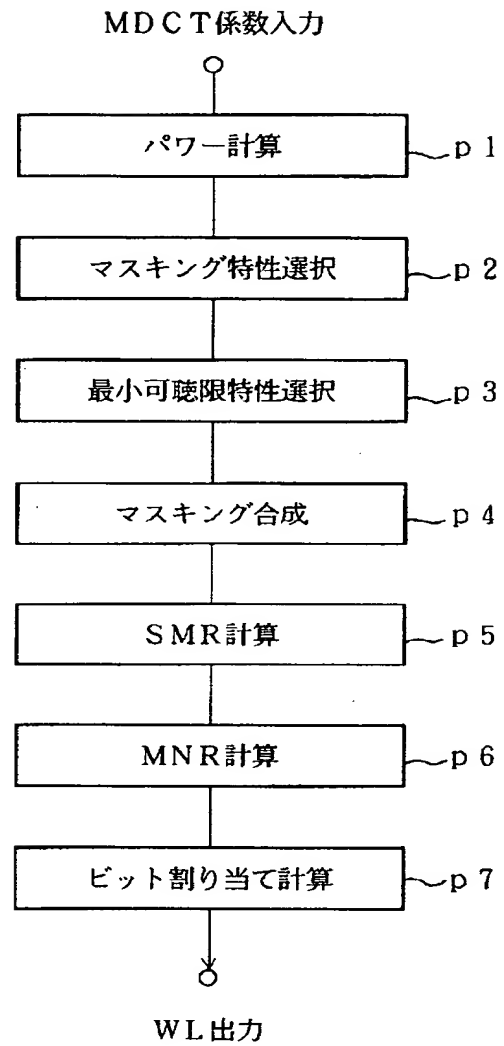
【符号の説明】

- | | |
|-------|-------------------|
| 1 | ミニディスク録音再生装置 |
| 4 | デジタルPLL回路 |
| 5 | 周波数変換回路 |
| 6 | 音声圧縮回路 |
| 6 a | テーブルROM |
| 30 7 | ショックプルーフメモリコントローラ |
| 8 | 信号処理回路 |
| 9 | ショックプルーフメモリ |
| 12 | 記録ヘッド |
| 13 | 光磁気ディスク |
| 21 | 光ピックアップ |
| 23 | 音声伸長回路 |
| 24 | デジタル／アナログ変換回路 |
| 36 | システムコントロールマイコン |
| 36 a | レジスタ |
| 40 37 | 入力操作手段 |

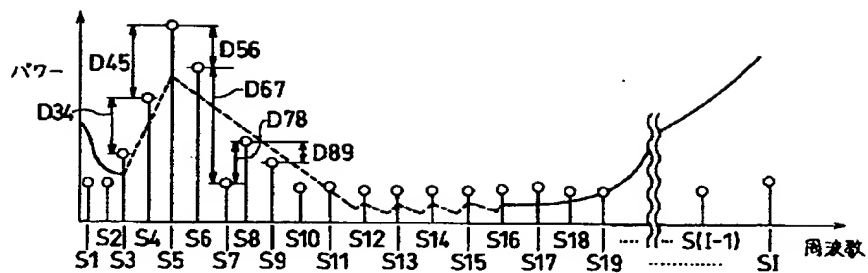
【図1】



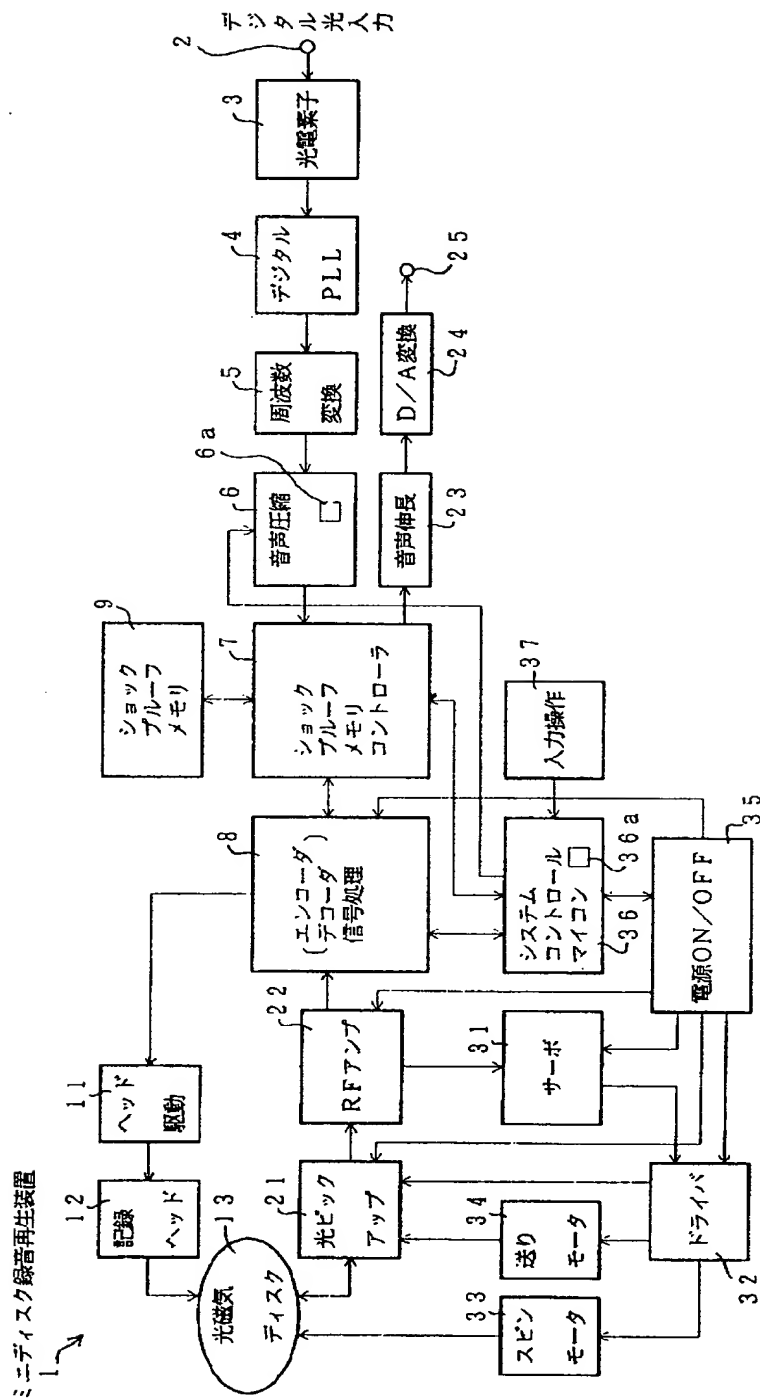
【図3】



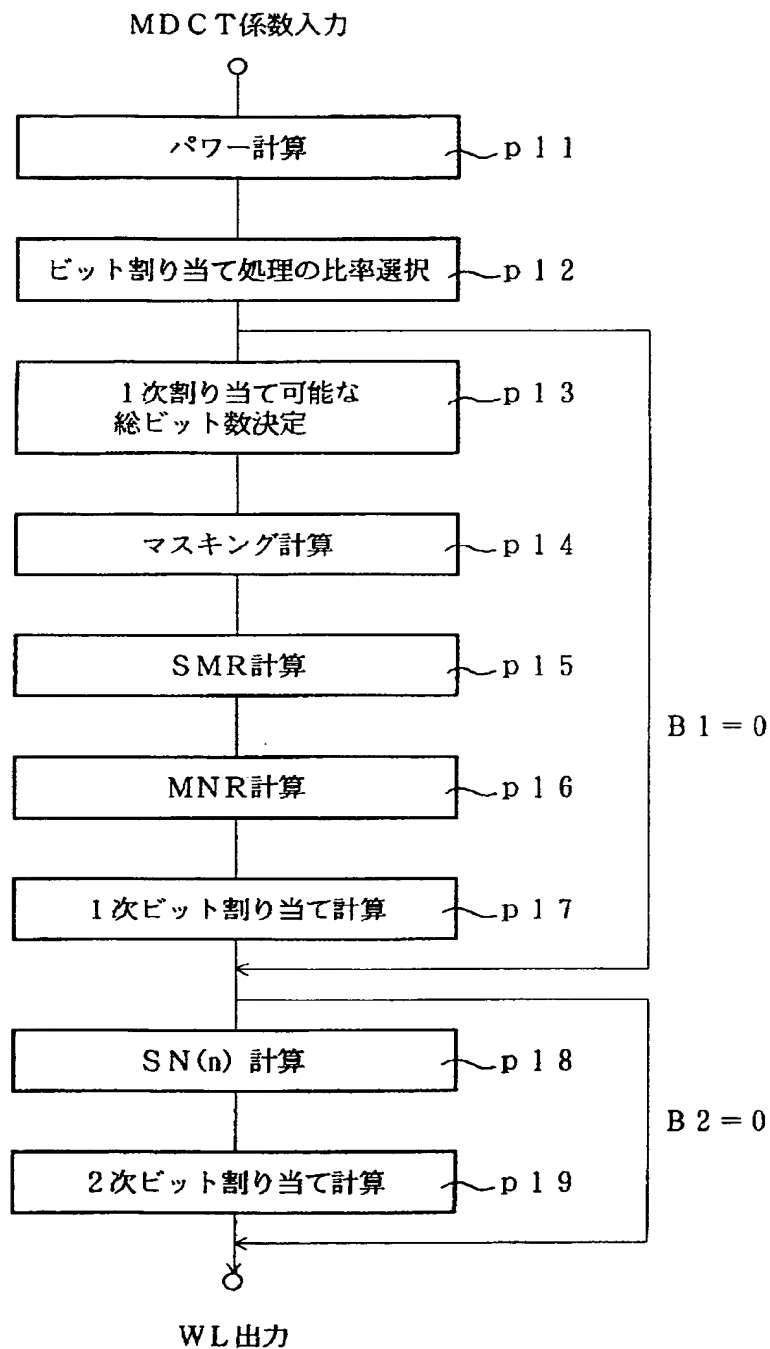
【図6】



【図2】



【図4】



【図5】

